Evaluating Technology-Based Strategies for Enhancing Motivation in Mathematics

Evaluating Technology-Based Strategies for Enhancing Motivation in Mathematics

Jon R. Star

Harvard University

Jason A. Chen

The College of William and Mary

Megan W. Taylor

Sonoma State University

Kelley Durkin

University of Louisville

Chris Dede

Harvard University

Theodore Chao

Ohio State University

To appear in *International Journal of STEM Education*

Accepted, September 19, 2014

Author Note

Correspondence should be sent to: Jon R. Star, 442 Gutman Library, 6 Appian Way, Harvard Graduate School of Education, Cambridge, Massachusetts, 02138; 617-496-2511 (voice), 339-368-7670 (fax), Email: jon_star@harvard.edu.

The research was supported by a grant from the National Science Foundation (DRL #0929575) to Chris Dede and Jon R. Star. The ideas in this paper are those of the authors and do not represent official positions of the National Science Foundation.

Thanks to Adam Seldow, Greg Jastrzemski, and the faculty, administration, and students of Chesterfield County Public Schools for their enthusiastic participation in the project. Thanks to Stephanie Fitzgerald for her assistance with all aspects of the project, and to Kinga Petrovai, Bharat Battu, Kevin Reeves, Arielle Niemeyer, Joy Casad, Chad Desharnais, Maisy Suslavich, Lauren Schiller, and Amy Venditta for their assistance with data collection and analysis.

Abstract

Background, context, and purpose of study: During the middle school years, students frequently show significant declines in motivation toward school in general and mathematics in particular. One way in which researchers have sought to spark students' interests and build their sense of competence in mathematics and in STEM more generally is through the use of technology. Yet evidence regarding the motivational effectiveness of this approach is mixed. Here we evaluate the impact of three brief technology-based activities on students' short-term motivation in math. 16,789 5th to 8th grade students and their teachers in one large school district were randomly assigned to three different technology-based activities, each representing a different framework for motivation and engagement and all designed around an exemplary lesson related to algebraic reasoning. We investigated the relationship between specific technology-based activities that embody various motivational constructs and students' engagement in mathematics and perceived competence in pursuing STEM careers.

Results: Results indicate that the effect of each technology activity on students' motivation was quite modest. No gains were found in self-efficacy; for implicit theory of ability, a lower incremental view of ability was found; we found modest declines in value beliefs. With respect to math learning, students in all three inductions had modest improvements in their scores on the math learning measure. However, these effects were modified by students' grade level and not by their demographic variables. In addition, teacher-level variables did not have an effect on student outcomes.

Conclusions. The present findings highlight the importance of tailoring motivational experiences to students' developmental level. Our results are also encouraging about developers' ability to create instructional interventions and professional development that can be effective when experienced by a wide range of students and teachers. Further research is needed to determine the degree, duration of, and type of instructional intervention necessary to substantially impact multi-dimensional, deep-rooted motivational constructs, such as self-efficacy.

Keywords: STEM education, technology, motivation, algebraic reasoning, self-efficacy, implicit theories of ability

Evaluating Technology-Based Strategies for Enhancing Motivation in Mathematics

Success in algebra during the middle grades is widely recognized to be a critical gatekeeper that constrains students' decisions about whether to pursue further educational opportunities in Science, Technology, Engineering, and Mathematics (STEM) fields (Adelman, 2006). Unfortunately, during this developmental period many students show significant declines in motivation toward school in general and mathematics in particular (e.g., Archembault, Eccles, & Vida, 2010; Blackwell, Trzniewski, & Dweck, 2007). One way that researchers have sought to spark students' interests and build their sense of competence in mathematics is through the use of various technological media. These technologies have ranged in complexity and cost from the simple and inexpensive, such as repurposing television programs, to the more complicated and expensive, such as specially designed mathematical experiences based on immersive virtual environments and computer games.

Despite the widely accepted notion that all technology-based activities are inherently engaging, the evidence regarding their motivational effectiveness is mixed (Moos & Marroquin, 2010). Part of the reason may be that many different types of technologies are available, and each can be designed well or poorly to leverage various aspects of motivation (e.g., engagement, self-efficacy, tenacity) in different ways. Another explanation for these mixed findings is that much of the research on technology-based activities considers motivation as a unidimensional construct intrinsically generated by technology usage rather than as a construct with multiple dimensions that may be impacted via various affordances of technology. This latter reason may be due to many developers lacking strong theoretical grounding in well-studied motivation constructs (Author, 2013; Moos & Marroquin, 2010).

As a step toward improving our understanding of the potential impact of technology-based activities on students' motivation in mathematics, the goal of this project was to investigate the relationship between (a) specific technology-based activities that exemplify various motivational constructs, (b) students' engagement in mathematics and perceived competence in pursuing STEM careers, and (c) students' mathematics learning from a short algebra lesson. As part of a four-day school-based intervention, students in grades 5 to 8 in a large school district were randomly assigned to three different technology-based activities, each representing a different framework for motivation and engagement designed around an exemplary lesson related to the learning of algebra.

Our research questions were as follows. First, what is the impact of the four-day intervention on students' motivation in mathematics, including interest in pursuing STEM careers? Second, to what extent is this impact influenced by factors such as the type of technological induction the students received and/or students' demographic and academic characteristics (e.g., gender, race/ethnicity, prior achievement)? Third, to what extent is this impact influenced by teacher-level factors such as credentialing in mathematics education, undergraduate major, years of experience, and teachers' beliefs (e.g., teaching self-efficacy)?

We begin by reviewing evidence on how and why technology-based activities might impact students' motivation in STEM fields.

Motivating Students to Learn STEM

As the National Academy of Sciences (2011) indicated, certain key ingredients are relevant for students who want to pursue STEM careers. These ingredients include a robust confidence in math and science capability, the ability to see one's abilities in STEM as able to improve over time, and the ability to develop a passion or sustained interest in becoming a scientist or engineer. Within the educational psychology literature, these key ingredients translate into three constructs, each of which has received substantial attention in the field of motivation: self-efficacy, implicit theories of ability, and value beliefs. We discuss each in turn.

Capable students plagued by a loss of confidence about their capacity to succeed in math and science typically avoid careers that require a strong background in those subjects (Lent et al., 2005). Decades of research have shown that students' self-efficacy, defined by Bandura (1997) as "the belief in one's capabilities to organize and execute courses of action required to produce given attainments" (p. 3), is a powerful influence on motivation and achievement. Bandura (1997) hypothesized several sources of self-efficacy, including *mastery experience* (the interpreted results of one's past performance), *vicarious experience* (observations of others' activities, particularly individuals perceived as similar to oneself), and *physiological and affective states* (anxiety, stress, and fatigue) – each of which has been linked to performance in math and science, including students' persistence in STEM fields and choice of STEM majors (Andrew, 1998; Author, 2010; Beghetto, 2007; Britner & Pajares, 2001; Gwilliam & Betz, 2001; Lau & Roeser, 2002; Lent, Brown, & Larkin, 1984; Luzzo, Hasper, Albert, Bibby, & Martinelli, 1999).

Accumulating evidence demonstrates that underrepresentation of women and racial/ethnic minorities may be substantially explained by considering the sources of self-efficacy. For example, Lent, Lopez, and Bieschke (1991) found that gender differences in math self-efficacy could be accounted for by students' mastery experiences, suggesting that women viewed their past experiences with math and science in a more negative light than did their male counterparts. Zeldin and Pajares (2000) found that women's decision to stay in the STEM pipeline could be attributed to the (vicarious) role models with whom they strongly identified, as well as the powerful social persuasions that came from women's most trusted sources (e.g., a mentor). Men, however, drew mostly from their mastery experiences—discussing their past successes and accolades as reasons for staying in the STEM pipeline. Therefore, in influencing students' participation in STEM fields, educators would be wise to look toward the sources that feed each individual student's self-efficacy to pursue such careers.

Like self-efficacy, implicit theory of ability (defined as a belief about the nature of intellectual ability (Dweck & Leggett, 1988)) plays an important role in motivation. Some individuals believe that their abilities are a fixed characteristic, and that nothing can be done to change that (i.e., “I’m not smart in math, and there isn’t anything I can do about it”). This is referred to as a *fixed theory* of ability. On the other hand, other individuals believe that, with sufficient effort and the proper strategies, one can become more able (i.e., “If I work hard in my math class, I can get smarter in math”). This is known as a incremental theory of ability. A large body of research has shown that implicit theory of ability plays a key role in students’ academic motivation, achievement, and career choices (Author, 2010, 2012; Blackwell, Trzesniewski, & Dweck, 2007; Cury et al., 2006; Good, Rattan, & Dweck, 2012; Grant & Dweck, 2003; Hong et al., 1999; Stipek & Gralinsky, 1996). For example, Blackwell, Trzesniewski, and Dweck (2007) found that, although Grade 7 math students’ who held a fixed theory of ability and those who held an incremental theory of ability both started at the same level of math achievement, by the end of the two years students who held an incremental view of ability achieved higher grades in math than did their fixed theory peers. Related, other work has suggested that teachers’ beliefs about the nature of intelligence may promote students’ conceptions of ability (Good, Rattan, & Dweck, 2012; Rattan, Good, & Dweck, 2012) and that gender and ethnicity may influence students’ conceptions of ability (Good, Aronson, & Inzlicht, 2003).

If, as Dweck and her colleagues have suggested, an incremental theory of ability can serve a protective function for students’ motivation and achievement, it would benefit researchers and educators to know what the sources of such a belief are. Little research has investigated this topic, however. Some studies suggest that process feedback highlighting the strategies and effort that lead to success can promote the view that ability is augmentable, whereas product feedback highlighting the accomplishments, but leaving out the perseverance required to get there, promotes a fixed view of ability (see Dweck & Master, 2009 for a review).

In addition to the self-efficacy and implicit theories of ability, value beliefs are also a significant determinant in students' motivation and achievement (Eccles et al, 1983). Value beliefs in mathematics and science deal with the question, "Do I want to pursue more opportunities in mathematics and science?" Eccles et al. defined values as being composed of several distinct constructs. First, students' *interest* or intrinsic value can affect the activities they pursue—activities that are more enjoyable are more likely to be pursued than are activities that are perceived to be lackluster. Second, students' perceptions of the *utility* of an activity refer to how valuable students perceive an activity to be. If an activity is perceived to be a steppingstone toward students' desired future endeavors, then students are more likely to pursue it. Finally, doing well in mathematics and science may influence students' identity or feelings of self-worth. This *attainment* value describes how important doing well in mathematics and science is to students' identity or feelings of self-worth.

Numerous studies have found that interest value predicts STEM career choice (Lent, Lopez, Lopez, & Sheu, 2008; Lent, Paixão, Silva, & Leitão, 2010), as well as choice in taking STEM courses (Eccles et al., 1984; Watt, Eccles, & Durik, 2006). Attainment value in mathematics and science is closely aligned with students' mathematics and science identity. The empirical literature supports that persistence and success in STEM careers may be rooted in students' identification with the roles and work of STEM professionals (Bonous-Hammarth, 2000; Estrada, Woodcock, Hernandez, & Schultz, 2011; Hernandez, Schultz, Estrada, Woodcock, & Chance, 2013). As such, attainment value predicts students' persistence in STEM careers (Carlone & Johnson, 2007; Oyserman & Destin, 2010).

Empirical literature also supports the notion that students' utility value predicts STEM success and choices. For example, Maltese and Tai (2011) found that students who perceived science to be useful were more likely to major in STEM subjects in college. Some have found convincing students that mathematics is useful for their future endeavors increased the interest of students only if they had high expectancies for success; those who expected to do poorly lost interest. However, Hulleman, Godes, Hendricks, and Harackiewicz (2010) found that, instead of telling students about the importance of an activity, if students discovered the usefulness of an activity on their own, the interest of those who had low expectations for success increased. For those whose expectancies for success were already high, no changes in interests were observed. Therefore, utility value can be influenced if students discover the utility of a subject on their own, with positive consequences for motivation and achievement.

Motivation and Technology

How can the constructs described above be targeted through technology-based educational experiences to support the motivation of students in mathematics and science? Although the literature on technology and motivation is quite large, relatively few of these studies employ frameworks that are grounded in well-studied psychological theories of motivation (Moos & Marroquin, 2010). Moos and Marroquin noted that the results about the effectiveness of technology as a motivational tool are mixed. One might expect lackluster outcomes if technology is applied as a "secret sauce" to automatically enhance students' engagement, rather than utilized in a principled manner to help an individual to find a robust sense of confidence in math and science capability, see his or her abilities in STEM as able to improve over time, and develop an interest for becoming a scientist or engineer.

With regard to self-efficacy, there is some evidence that engagement with innovative technology in academic settings can positively impact self-efficacy toward STEM. For example, Ketelhut and colleagues (Ketelhut, 2007; Ketelhut, Nelson, Clarke, & Dede, 2010) found that students' self-efficacy for scientific inquiry before using a Multi-User Virtual Environment (MUVE) called River City was related to their behaviors within the virtual world. In particular, less self-efficacious students manifested a self-efficacy boost through mastery experiences gained through engagement in the activities of the MUVE. Similarly, Liu, Hsieh, Cho, & Schallert (2006) explored middle school students' science learning within a computer-enhanced Problem Based Learning (PBL) environment called *Alien Rescue* and found that students' achievement and self-efficacy increased after participating in *Alien Rescue*.

Building on studies such as these, one additional promising avenue in exploring how innovative technologies can be used to tap the sources of self-efficacy deals with the capability to use virtual representations of the self (avatars) in creative ways. For example, Fox and Bailenson (2009) reported that, when individuals watched a virtual representation of themselves experiencing the benefits of exercising, these individuals were significantly more likely to engage in exercise after the intervention was done. In contrast, individuals who watched virtual representations of themselves loitering did not engage in exercise after the intervention nor did individuals who watched a virtual representation of others. As another example, Rosenberg-Kima, Baylor, Plant, and Doerr (2008) reported greater gains in self-efficacy for pursuing engineering careers when participants saw virtual avatars on a computer interface who looked similar to themselves. These results suggest that virtual models of a person successfully attempting a task can be effective in shaping a person's self-efficacy and behavior.

Technology also seems to be a promising avenue for impacting implicit theory of ability. In particular, Dweck and her colleagues have developed a web-enabled intervention, Brainology®, based on the paper and pencil version of their curriculum materials designed to enhance implicit theory of ability. Students are introduced to two cartoon characters who guide them through the web-based environment, where they learn about the functions of the brain, including that the brain is like a muscle—with conditioning, it can get stronger – an attitude which is linked to an incremental view. Donohoe, Topping, and Hannah (2012) conducted a quasi-experimental study on 33 adolescents (ages 13-14) and found that Brainology® led to a significant increase in students' incremental view of ability. More generally, although a substantial literature base has shown that manipulating students' beliefs about the plasticity of ability leads to positive motivational and achievement gains, the research base concerning how technologies can be used to tap this construct is quite limited.

With respect to value beliefs, the research base about technology is similarly small. However, researchers have argued that well-designed technology-based activities can be used to target students' interest value beliefs by making learning goals relevant and meaningful, and by allowing students to identify with characters within the technology environment (Gee, 2003; Squire, 2003). For example, Moos and Azevedo (2008) found that a hypermedia environment enhanced the development of students' interest but not their utility value beliefs. Similarly, Hickey, Moore, and Pellegrino (2001) showed that the use of *The Adventures of Jasper Woodbury* videodisc activity led to gains in students' mathematics interest, although these gains appeared to result both from the technology as well as from teachers' beliefs and instructional practices.

Context of the Present Study

To investigate the potential impact of technology-based activities on students' mathematics motivation, we designed three different types of technology activities (or 'inductions'). (We use the term 'induction' to refer to the technology activities, to avoid possible confusion between the technology activities and math lesson activities (described below).) The inductions differed along two main dimensions. First, the design of each induction was based on a different motivational construct; in other words, the theory of change underlying each induction differed (as we elaborate below). Second, the inductions differed in the expense and technical sophistication that were required for their creation and implementation, ranging from the very expensive-to-produce and technically advanced to the modest and inexpensive. Below we describe each induction in more depth.

Induction 1: Virtual Environment

At the core of Induction 1 was an Immersive Virtual Environment (IVE) - a game-like activity we designed to introduce students to the mathematical concepts that were to follow in a subsequent lesson. The IVE was professionally produced such that it was similar in look and feel to video games that students may have had experience playing.

For the storyline of the IVE, students were provided with the opportunity to explore an outer space environment in the context of a space rescue mission. Various mathematical puzzles were encountered as students moved around the planet; all puzzles related to the generation of and identification of mathematical patterns, similar to what would subsequently be discussed in a mathematics lesson. The initial puzzle was designed to be relatively easy; in later stages of the experience, mathematically-related, more complex puzzles were broken down into many smaller steps to scaffold students' progress and to reduce the likelihood that students would be overly frustrated. Similarly, hints were also provided by the IVE for students who requested help in completing any of the puzzles.

Prior to beginning the IVE, each student viewed a short (5-minute) video clip of a young STEM professional who talked about the nature of the work they do (e.g., designing astronaut space suits), the difficulties they had encountered in their K-12 math and science classes, and how they were able to overcome these difficulties. Students were provided with a selection of several of these videos, which varied according to the demographic attributes of the STEM professionals (e.g., gender, ethnicity); students were allowed to select whichever single video they wanted to view before beginning the IVE.

Motivationally, Induction 1 was designed to primarily impact students' self-efficacy. In particular, we attended to the sources of self-efficacy beliefs theorized by Bandura (1986, 1997) and described above. The IVE experience supported mastery experiences by allowing students to experience incrementally more difficult mathematical challenges, and by providing the scaffolds necessary for students to succeed when they were met with obstacles. Vicarious experiences were included in Induction 1 by including real-life, young, STEM professionals who discussed their jobs and the types of obstacles that they faced (and overcame) as they pursued a STEM career. Finally, emotional and physiological states were addressed by ensuring that students felt comfortable and relaxed about solving the mathematical challenges in the IVE. For example, we made the design decision *not* to include a timer that gently reminded students to work more quickly if they were taking too long, because such a timer would likely cause a good deal of anxiety—a common experience for many students in mathematics.

Induction 2: Brainology® Web-based Activity

For the second induction, we used a commercially-available series of web-based modules designed to teach students about an incremental view of ability. These modules are based on the work of Dweck and colleagues and have been shown to be successful at influencing students' motivation and achievement (e.g., Blackwell, Trzesniewski, & Dweck, 2007). Students assigned to Induction 2 were given access to an abridged version of the Mindset Works® StudentKit - Brainology® program (www.mindsetworks.com). (This abridged version was created by Dweck and colleagues specifically for the present study.) In a series of interactive modules, animated characters taught students how the brain works and how the brain grows stronger with effort. Students progressed through the modules at their own pace. The intervention that students experienced was relatively short compared to the entire Brainology® program, which contains over two hours of online instruction and up to 10 hours of additional activities to do over a recommended period of 5 to 16 weeks. Brainology is specifically designed for 5th to 9th grade classrooms. Note that the Brainology® modules do not have a specific focus on mathematics, nor do they incorporate any mathematical problem solving or algebraic reasoning.

With respect to motivation, the Brainology® program is explicitly designed to impact students' implicit theory of ability. As noted above, Dweck and her colleagues (Dweck & Leggett, 1988; Blackwell, Trzesniewski, & Dweck, 2007) have shown students possess particular 'mindsets' that can influence their motivational and developmental trajectories through the course of school (e.g., fixed theory of ability vs. incremental theory of ability). The Brainology® program activities have been found to encourage students toward a incremental view of ability.

Induction 3: Video on Mathematical Patterns

Induction 3 was intended to provide an off-the-shelf experience for students related to some of the mathematical ideas that were to come in the mathematics lesson. We selected a commercially available PBS NOVA video on fractals because of its engaging storyline and graphics, its focus on mathematical patterns, and the accessibility of the content to our target population of students in grades 5-8. The 2009 video, *Fractals: Hunting the Hidden Dimension*, is 56 minutes long and includes visually appealing animations, interviews with mathematicians, and accessible explanations of the mathematics of fractals and their applications to everyday life, such as building smartphone antennas and generating visual effects in movies.

In terms of motivation, movies have long been used by educators to motivate and engage students in the classroom. Although this movie did not specifically target a particular motivation construct, movies are often used in educational settings as an inexpensive, simple means that teachers can employ to help students see connections between what they are learning and real-world applications.

Mathematics Content Focus

Within the general landscape of STEM, we chose to situate the present study in the content area of algebra. Algebra is widely recognized as a crucial peg in the trajectory of mathematical learning, because of the conceptual and procedural groundwork it lays for accessing higher mathematics and because it presents a shift in how students are expected to think mathematically (Kieran, 1992). Algebra is often the first time students are introduced to some of the most important and useful ideas in the field of mathematics, such as the concept of a “variable” or the generalization of patterns in generated data (Author, 2009). However, students’ difficulties in algebra are well-documented on both national and international assessments (e.g., Beaton et al., 1996; Blume & Heckman, 1997; Lindquist, 1989; Schmidt, McKnight, Cogan, Jakwerth, & Houang, 1999). For example, in the eighth-grade data from the US National Assessment of Educational Progress [NAEP] show that students continue to struggle on very straightforward algebra problems: Only 59% of 8th graders were able to find an equation that is equivalent to $n + 18 = 23$, and only 31% of 8th graders were able to find an equation of a line that passes through a given point and with a negative slope (National Assessment of Educational Progress, 2011).

Within the larger landscape of algebra, we focus here on an aspect of algebra that many mathematics educators refer to as algebraic reasoning (e.g., Kaput, 1999), which includes using arithmetic for generalizing, working with patterns to describe functional relationships, and modeling as a way to formalizing generalizations. Algebraic reasoning has begun to play an increasingly important role in US mathematics instruction, as evidenced by its emphasis in several grade levels of the Common Core Standards (Common Core State Standards Initiative, 2010). Furthermore, the exploration and modeling of data that lie at the core of algebraic reasoning are central to the work of scientists, engineers, and other STEM professionals (Hoyles, Noss, Kent, & Bakker, 2010). In many middle grades mathematics classrooms, algebraic reasoning is instantiated through the identification, justification, and generalization of numerical patterns in given or generated data.

At the core of the present study is a two lesson mathematics activity in which students engaged in an exploration of mathematical patterns. We designed the activity around a combinatorics task often referred to as a "trains" problem, because it involves the creation of integer-length "trains" using different numbers and lengths of integer-length "cars." For example, students may be asked to determine the number of possible trains of a certain length *n* that can be created. If the task is to create a train of length 4, there are 8 ways to do so (using only integer-length cars, where the order of the cars matters): 1-1-1-1, 1-1-2, 1-2-1, 2-1-1, 1-3, 3-1, 2-2, and 4. Similarly, to make a train of length 5, there are 16 ways to do so. There are a large number of interesting variations and extensions of the trains problem, such as: How many trains of length *n* can be made using only cars of length 1 and 2, or only with cars of length 2 and 3? Or how many trains of length *n* can be made that begin with a car of a given length?

The trains problem was a useful context in which to ground our study for the following reasons. First, the mathematical content of the trains problem, which includes identifying, justifying, and generalizing numerical patterns, is well-aligned with current state and national content standards for algebra. Second and similarly, the instructional practices involved in optimally implementing the trains problem (including use of mathematical manipulatives or representations to depict the trains, small group work leading to whole class discussions, and the sharing and comparing of students' problem solving strategies) are consistent with current "best practices" in mathematics education (e.g., Common Core State Standards Initiative, 2010, National Council of Teachers of Mathematics, 2000). Third, as noted above, the intellectual activities of the trains problem, including exploring and modeling data, are central to the work of many STEM professionals. And finally, the trains problem is approachable to students from a variety of grade levels.

An overview of the math activity is as follows (see Figure 1). The lesson was designed to occur on two consecutive days; teachers were given latitude to decide where the break between the first and second days of the lesson would occur. Teachers were provided with a variety of materials to aid in their implementation of the lesson, including detailed and condensed lesson plans, poster-sized visual aids, and concrete and virtual manipulatives.

The Current Study

The goal of the present study was to investigate the relationship between specific technology-based activities and students' motivation in math. Students in grades 5 to 8 participated in a four-day classroom-based experience, beginning with a one-day technology activity, followed by a two-day mathematics lesson on algebraic reasoning, and concluding with revisiting the same technology induction on the final day. Students were assigned to one of three different types of technology inductions (as described above), each representing a different motivational framework. An assessment that targeted motivation was administered before, immediately after, and two months after the intervention.

We hypothesized that Inductions 1 and 2 would have the strongest effect on the motivational constructs that they were designed to influence. In particular, we hypothesized that Induction 1 would have the strongest impact on students' self-efficacy and that Induction 2 would have the strongest impact on students' implicit theory of math ability. Because Induction 3 was not designed with a particular theory of motivation in mind, it did not intentionally target any particular motivation variable. However, because of the content in the movie, we hypothesized that this third induction would have an impact on students' value beliefs, especially their utility and interest value. Finally, with respect to developmental issues in motivation, the literature is clear that there is a general decline in motivation as students progress through school (Archambault, Eccles, & Vida, 2010; Eccles, Midgley, & Adler, 1984). Because the structure of schooling for students in middle school (Grades 6-8) is different from that of elementary school students (Grade 5), and because students conceive of competence differently based on age (Dweck, 1986), we expected the first two inductions to have differential impacts on students depending on their age.

Method

Sample

Data come from all 5th, 6th, 7th, and 8th grade students and their teachers in the Chesterfield County Public School district in Virginia. A total of 18,628 students participated in the study, along with their 476 teachers, from 38 elementary and 12 middle schools.

A number of teachers in our original teacher pool were assistant, ESL, or special education teachers who did not have their own classroom. We removed these teachers from our sample, ending up with 339 teachers in our active teacher sample who participated in random assignment. In the elementary schools, the 163 5th grade teachers, who taught all subjects to the same group of students each day, implemented the intervention with their homeroom students. In the middle schools, the 60 6th, 57 7th, and 59 8th grade teachers were all math specialists and implemented the intervention in each mathematics classes that they taught. In total, the intervention was implemented in a total of 545 distinct mathematics classes.

We removed students who did not have parental consent to be a part of the study, which left us with 16,879 students. In addition, we had to exclude the 8,979 students (and their 113 teachers from 5 schools) who were missing pretest or posttest data used in our analyses. Most of this missing data was due to a miscommunication between the research team and the district relating to the student identification numbers that students were instructed to use at pre-test. Almost 5,000 students used an incorrect identification number, making it impossible to match students' pre- and posttest scores. Little's (1988) Missing Completely at Random (MCAR) test confirmed that these data were not missing completely at random (χ^2 (1576)= 7162.88, $p < .001$). In particular, students with missing data were more likely to be male, African-American or Hispanic/Latino, with ELL status, and from schools with a high percentage of free or reduced lunch. After removing those students with missing data, we report on the 7,900 students and 226 teachers from 44 schools who remained in our analyses.

Due to the large amount of missing data (about 53% of students were missing demographic, pre- or posttest data), it was not advisable to use multiple imputations to include more of these students and teachers in our analyses. As a result, we compared those participants included to those excluded using χ^2 tests and t-tests to examine differences in our demographic and pretest variables. We found several differences (see Appendix Table A-1). For instance, excluded participants were more likely to be male, African-American or Hispanic/Latino, and to have ELL status. They were also more likely to come from schools with a higher percentage of students receiving free or reduced lunch. There were few significant differences between the groups on student pretest variables, with the one exception being that excluded students had lower self-efficacy than included students (p = .037). There were significant differences between the groups on several teacher pretest variables. The excluded participants had teachers with lower self-efficacy for student engagement and instruction (p = .002) and self-efficacy for technology use ($p < .001$) than included participants. However, excluded participants had teachers with higher mathematics self-efficacy ($p < .001$) than included participants. The implications of these differences are examined in the discussion section.

The included 7,900 students were approximately equally divided across grade levels (see Table 1). The majority of students (60%) were White, with 23% African-American, 8% Hispanic, and 3% Asian. Four percent of students were identified as English-language learners [ELLs]. School level information was collected about students' eligibility for free or reduced lunch; participating schools had an average of 34% of students who were eligible for free or reduced lunch, with eligibility at the school level ranging from 2% to 85%. We also collected students' most recent scores on the state standardized test in mathematics, the Virginia Standards of Learning (VA-SOL) test; this test is given annually to students in grades 3-8.

Design and Procedure

We used a pre-test/post-test[1] experimental design. Prior to the start of the intervention, students and teachers were administered a pretest. After pre-test administration, teachers were randomly assigned to one of three inductions described above – see Table 1 for student demographics for each induction. Participation in the main part of the intervention occurred over a period of four consecutive days. On Day 1, students worked on the induction to which they were assigned. On Days 2 and 3, teachers taught the two-day mathematics lesson. On Day 4, students again worked on the induction to which they were assigned.

For students in Induction 1, Day 1 of the intervention was spent in the school's computer lab. Each student sat at his/her own computer, with headphones, and watched the short interview of a STEM professional and then played the IVE game for approximately 30 minutes. On Day 4, students returned to the computer lab and restarted the technology-based activity, including watching a video of a STEM professional and restarting the IVE game from the beginning – again playing for about 30 minutes. Similarly, for students in Induction 2, Days 1 and 4 were spent in the school's computer lab, with one student at each computer with headphones, playing the Brainology® program. Finally, Induction 3 students watched the first half of the *Fractals: Hunting* *the Hidden Dimension* video (about 28 minutes) on Day 1; on Day 4, these students watched the second half of the video.

Professional Development

All teachers were provided with a one-day (6.5 hours) professional development (PD) workshop, administered within one week of the start of the intervention. The PD workshop was designed and implemented by project staff. An identical PD was repeated for five consecutive days; district administration determined which teachers would attend on each day, with the attendance ranging from 56 teachers to 123 teachers. Each PD workshop included teachers from all three inductions and all four grade levels.

Most of the PD (approximately 4 hours) was devoted to introducing teachers to the two-day mathematics lesson. Teachers were provided with detailed lesson plans as well as visual aids, handouts, and manipulatives that accompanied the lesson. Under the facilitation of the first author, an experienced mathematics teacher educator, the PD workshop provided teachers with the opportunity to engage with the mathematics of the lesson and to plan for the enactment of the lesson. Approximately one hour of the PD was spent providing teachers with an overview of the project procedures, measures, and logistics. For the remainder of the PD, we provided teachers with induction-specific training. Teachers were divided into groups based on which induction they were assigned to. Induction 1 teachers played the IVE in a computer lab, Induction 2 teachers explored the Brainology® program in a different computer lab, and Induction 3 teachers watched the *Fractals: Hunting the Hidden Dimension* movie in a seminar room.

Measures

All assessments were administered to teachers and students online, via a password-protected website.

Student motivational measures. All students were administered a pre- and post-assessment, in a proctored computer lab in each school, during the regular school day. The pre-test, taken between one and three weeks prior to the start of the intervention, targeted students' motivation, with measures corresponding to the three motivational constructs that were related to the inductions – self-efficacy, implicit theories of ability, and value (see Table 2 for descriptive information on student variables; see Table 3 for sample items and alphas).[2] The post-test was administered on Day 4, after the implementation was completed. The motivational items on the post-test were identical to the pre-test. As described below, we used validated scales that have been commonly used in other motivation studies. Also, an exploratory factor analysis and scree plot indicated that our items mapped well onto three factors, with all self-efficacy items loading best onto one factor (factor loadings from 0.59 to 0.72), all value items loading best onto the second factor (factor loadings from 0.41 to 0.71), and all implicit theories of abilities items loading best onto the third factor (factor loadings from 0.53 to 0.61).

We assessed self-efficacy students with a 13-item measure that was drawn from Bandura's (2006). The degree to which students endorsed an incremental view of ability (as opposed to a fixed view of ability) was assessed using a 6-item instrument that was adapted from Dweck (1999). Note that for the analysis of implicit theory, we reverse-scored the fixed theory of ability items and calculated a mean theory of ability score with the incremental items – thus higher scores represented stronger agreement with incremental theory of ability. Finally, interest, attainment, and utility value beliefs concerning their mathematics class were assessed using scales taken from the Michigan Study on Adolescent Life Transitions (MSALT), which has been used extensively in the past (e.g., Eccles, Barber, Stone, & Hunt, 2003).

Student mathematics learning measure. Assessing students' mathematics learning was not a major focus of the present study, mainly because of the absence of *a priori* hypotheses related to the differential impact of the three technology inductions on student learning and also the short duration of the math lesson. However, as a manipulative check, we included a short five-item assessment on mathematics learning on both the pre- and post-tests. These five items were on algebraic reasoning as related to the two-day mathematics lesson, specifically data organization, pattern identification, and the ability to make generalizations. For example, an item on pattern identification asked students to identify the number that is most likely to come next in the number pattern: 3, 7, 11, 15, ?. As another example, an item asked students to determine how many different lunch plates could be made by choosing one main course (from two choices), one side (from four choices) and one drink (from two choices). Items on the pre- and post-tests were non-identical but isomorphic (same problem structure but with different contexts and numbers). The reliability of the math learning measure was low ($\alpha = 0.30$ and 0.40 for the pre- and post-test); as a consequence, the results from this measure must be interpreted with caution.

Teacher measures. All teachers were administered three assessments. Teachers completed the surveys at a time (within a given survey administration window) and place of their choosing.

First, teachers were given a pre-test immediately prior to the start of the professional development workshop. The pre-test collected background and demographic information about teachers, such as number of years teaching, undergraduate major, advanced degrees held, and national board certification status. In addition, the teacher pre-test included items that tapped teachers' own teaching self-efficacy for instruction and student engagement (22 items), technology use (7 items), and mathematics (12 items). Items were drawn or adapted from Bandura (2006). To confirm the validity of the self-efficacy items, we first conducted an exploratory factor analysis. This analysis indicated that our self-efficacy items mapped well onto three factors, with all self-efficacy items related to student engagement and instruction loading best onto one factor (factor loadings from 0.51 to 0.76), all self-efficacy items related to technology use loading best onto the second factor (factor loadings from 0.45 to 0.82), and all self-efficacy items related to mathematics loading best onto the third factor (factor loadings from 0.45 to 0.80). Teachers were also administered a 6-item measure of implicit theory of ability that was adapted from Dweck (1999). See Table 3 for sample items and alphas.

Second, teachers completed a 6-item post-professional development survey immediately after the one-day professional development workshop (see Appendix Table A-2). This survey assessed teachers' views on the overall quality of the professional development workshop, how prepared and confident teachers felt in implementing the intervention, and teachers' predictions about how students would react to this intervention. Finally, immediately after they had finished teaching the two-day math lesson, teachers were administered a six-item self-assessment of implementation fidelity asking about their adherence of this lesson plan.

Data Analysis

Given that many students had the same teacher and many teachers were in the same school, we used multilevel modeling (Raudenbush & Bryk, 2002) to account for this nesting of students within teachers and teachers within schools. The first level of the model, the student level, included students' prior knowledge (VA-SOL) scores, pretest math learning scores, pretest self-efficacy scores, pretest implicit theory of ability scores, pretest value scores, and demographic information, including ELL status, grade, gender (male coded as 1 and female coded as 0), and ethnicity.

The second level of the model, the teacher level, measured the effect of experimental condition, teachers' self-efficacy for student engagement and instruction, teachers' self-efficacy for technology use, teachers' mathematics self-efficacy, and teachers' implicit theory of math ability. We specified Induction 1 (the immersive virtual environment) as the referent condition to compare it to the other two inductions. This resulted in the effect of condition being captured by two variables. One variable indicated the difference between Induction 1 and Induction 2, and the other variable indicated the difference between Induction 1 and Induction 3. To test the difference between Inductions 2 and 3, a Wald test (similar to an incremental F test) was used to examine whether the parameter estimates for these conditions were significantly different from one another.

The third level of the model, the school level, measured the percentage of students receiving free or reduced lunch in each school. Finally, we also included two cross-level interactions to test for possible interactions between induction and grade, as well as two cross-level interactions to test for possible interactions between induction and prior math knowledge (VA-SOL). All continuous independent variables in the model were grand mean centered. We ran these models to evaluate our four posttest student outcomes: math learning, self efficacy, implicit theory of ability, and value.

The intraclass correlations for the teacher and school levels ranged from 0.001 to 0.052, which were fairly small. However, we still used multilevel models because they account for dependency between observations, and produce unbiased standard errors and more stable intercept and slope estimates (Myers, 2011). Similar results were obtained when using Ordinary Least Scales [OLS] regression instead of multilevel models.

Results

We begin by providing descriptive information on the quality of the implementation of the professional development workshop and the intervention, as well as by describing teachers' view of the quality of professional development, teachers' assessment of students' interest and engagement with the intervention, and teachers' self-reports of their fidelity of implementation. We then turn to our research questions by overviewing students' scores on the motivational variables at pretest and posttest and then reporting the effects of condition at posttest.

Fidelity of Implementation

Quality of the professional development. Judging from teachers' self-reported responses to the survey administered immediately after the PD (see Appendix Table A-2), teachers were not especially enthusiastic about the quality of the PD, with only 29% rating the experience as very good or excellent as compared to other PD experienced in the past five years. Nevertheless, a plurality of teachers left the PD feeling prepared to implement the math lessons (45% felt prepared or very prepared), and most felt confident that they could successfully do so (58% felt confident or very confident), despite the fact that very few teachers felt that the lesson was similar or very similar to the ways that they typically taught. Most teachers (53%) felt that students would be very challenged by the content of the math lessons, and many teachers (47%) felt that students would react positively.

Implementation of math lessons. Recall that data on fidelity of implementation were obtained from self-reports of teachers on the survey administered immediately after the end of the two-day math lesson. Teachers' responses indicated that they believed that they had very closely followed the lesson plan, with 75% indicating that they very closely or exactly followed the list of activities and 60% answering that they asked the questions very closely or exactly as suggested (see Appendix Table A-2).

Student and Teacher Pretest Scores

To begin, we measured whether there were any differences between the inductions on our outcome measures at pretest and on demographic variables (see Table 2). When controlling for other independent variables in the model, there were no significant differences ($p > .05$) between inductions on any of the pretest or demographic variables, with the exception of prior knowledge (VA-SOL). Students in Induction 2 had lower prior knowledge than students in Induction 1, $\beta = -15.76$, $p = .003$, and Induction 3, $\chi^2(2) = 13.63$, $p = .001$. Students in Induction 3 also had slightly lower prior knowledge than students in Induction 1, $\beta = -15.69$, $p = .001$. Prior knowledge was included in all subsequent models, so we controlled for these differences between conditions.

Pre/Post Gains

Before examining the effects of condition, we first consider whether the intervention generally led to gains in students' motivation (see Table 2). Overall, students did not have statistically significant gains on our measure of self-efficacy ($M_{pre} = 4.54$, $M_{post} = 4.55$, $t = -1.16$, $p = .246$, $d = -0.01$). For implicit theory of ability, students' incremental view of math ability decreased after the intervention, although this was a small effect ($M_{pre} = 4.22$, $M_{post} = 4.16$, $t = -6.93$, $p < .001$, $d = -0.07$). For value, students' scores generally decreased after the intervention as well, although the effect was again small ($M_{pre} = 4.24$, $M_{post} = 4.19$, $t = -8.71$, $p < .001$, $d = -0.06$). For math learning, the intervention led to an average gain on students' scores on the five-item mathematics learning assessment of ten percentage points, and this was a moderate effect ($M_{pre} = 0.60$, $M_{post} = 0.70$, $t = 28.60$, $p < .001$, $d = 0.40$).

Effects of Condition at Posttest

We now move to examining the effects of condition at posttest. At posttest, there were significant effects of condition on several of our outcome variables (see Table 4). As we describe below and return to in the discussion, note that most of the independent variables that significantly predicted our outcomes were at the student-level, rather than at the teacher-level.

For each analysis of the effect of condition, we report three interrelated analyses, in the following order. First, we report whether Induction 2 differed from Induction 1 (main effects and interactions), and we then report whether Induction 3 differed from Induction 1 (main effects and interactions). Finally, we report results from a Wald test to investigate whether Inductions 2 and 3 differed (main effects and interactions).

Math learning. Comparing Inductions 1 and 2, students in Induction 2 earned similar math learning scores to students in Induction 1, $\beta = 0.003$, $p = .872$. There was also no significant interaction between Induction 2 and grade, $\beta = 0.01$, $p = .129$. Comparing Inductions 1 and 3, students in Induction 3 had similar math learning scores to students in Induction 1, $\beta = -0.01$, $p = .409$. However, there was a significant interaction between Induction 3 and grade. In particular, students in lower grades benefited more from Induction 1 than from Induction 3. Then as grade increased, Induction 3 became more effective, $\beta = 0.02$, $p = .013$. Thus, for students in grade 5, being in Induction 1 led to higher scores on average. For students in grades 6, 7, and 8, being in Induction 3 led to higher scores on average. Finally, post-hoc Wald tests comparing Inductions 2 and 3 suggested that there were no significant differences between Inductions 2 and 3 ($\chi^2(2) = 1.06$, $p = .589$); however, there was a significant interaction when considering grade ($\chi^2(2) = 6.22$, $p = .045$). Essentially, Induction 2 was more effective for lower grades, and as grade increased, Induction 3 became more effective. There were no significant interactions between induction and prior knowledge (VA-SOL) (*p*'s > .532).

Self-efficacy. There were no significant differences between any of the inductions on the student self-efficacy variable, nor were there any significant interactions between inductions and grade or inductions and prior knowledge (*p*'s > .128).

Implicit theory of ability. Comparing Inductions 1 and 2, students in Induction 2 had higher implicit view of math ability scores than students in Induction 1, $\beta = 0.09$, $p = .039$, meaning that being in Induction 2 led to an implicit theory of math ability score that was 0.09 standard deviations higher than being in Induction 1. There was also a significant interaction between Induction 2 and grade. In particular, students in lower grades had similar implicit view of math ability scores in Induction 2 and Induction 1. Then as grade increased, Induction 2 led to higher implicit view of math ability scores than Induction 1, $\beta = 0.12$, $p < .001$. In addition, there was a significant interaction between Induction 2 and prior knowledge (VA-SOL), β = 0.001, p = .018; however, as the coefficient indicates, this was a very small interaction. Students with lower prior knowledge had slightly higher implicit view of math ability scores in Induction 1 than Induction 2. Comparing Inductions 1 and 3, students in Induction 3 had similar scores to students in Induction 1, β = 0.05, p = .243. There was also not a significant interaction between Induction 3 and grade, β = 0.03, p = .271, nor between Induction 3 and prior knowledge (VA-SOL), β < 0.001, p = .371. A post-hoc Wald test indicated that overall students in Induction 3 had similar implicit theory of ability scores to those in Induction 2 ($\chi^2(2) = 4.34$, p = .114. However, there was a significant interaction when considering grade ($\chi^2(2) = 23.62$, p < .001). In lower grades, students in Induction 3 had similar implicit view of math ability scores as students in Induction 2, but as grade increased, students in Induction 2 tended to have higher scores than students in Induction 3. When comparing Inductions 2 and 3, there was also a marginally significant interaction between Induction and prior knowledge (VA-SOL) ($\chi^2(2) = 5.75$, p = .057).

Value. For value, in comparing Inductions 1 and 2, overall students in Induction 2 had similar value scores to students in Induction 1, β = 0.02, p = .668. There was also no significant interaction between Induction 2 and grade, β = -0.01, p = .520. When comparing Inductions 1 and 3, students in Induction 3 had similar value scores to students in Induction 1, β = 0.01, p = .795. There was a significant interaction between Induction 3 and grade. In particular, students in lower grades had similar value scores in Induction 3 and Induction 1. Then as grade increased, Induction 1 led to higher value scores, β = -0.04, p = .036. Post-hoc Wald tests suggested that there was no significant difference between Inductions 2 and 3 ($\chi^2(2) = 0.19$, p = .910). There was also no significant interaction when considering grade ($\chi^2(2) = 4.76$, p = .093). Finally, there were no significant interactions between condition and prior knowledge (VA-SOL) (p's > .069).

Discussion

Perhaps not surprisingly given the size and complexity of the present study, our results are informative, modest, and not definitive. We begin by summarizing the results that pertain to our three research questions in turn, with particular attention to the contributions of these results to the field.

RQ1: Impact on Students' Motivation

Our first research question concerned the general impact of the four-day intervention on students' motivation in mathematics, particularly self-efficacy, implicit theory of ability, and value. Overall, results from the four-day intervention were mixed. No gains were found in self-efficacy; for implicit theory of ability, a lower incremental view of ability was found; we found modest declines in value beliefs. With respect to math learning, students in all three inductions had modest improvements in their scores on the math learning measure.

RQ2: Influences of Induction Type and Student Characteristics

Second, we were interested in whether the impact of the intervention was influenced by the type of induction that student received and other student-level demographic or academic characteristics. We found that induction type and student-level factors had a moderate influence on the motivational impact of the intervention. No effects related to self-efficacy were found, and effects related to value were very minor. For implicit theory of ability, there were indications that Induction 2 was more successful than Inductions 1 and 3 in impacting students' views, especially for older students. Induction 2 led to higher incremental views of math ability for students, particularly for students in grades 7 and 8. Induction type also appeared to have a small impact on value, with some evidence that Induction 3 had the strongest impact on utility and attainment value for the younger students, as compared to the other two inductions.

Despite the complexity of these results for our second research question, three clear patterns did emerge.

Absence of effects on self-efficacy. First, Induction 1 did not have the hypothesized impact on students' self-efficacy. Despite the fact that the IVE was designed specifically to foster changes in self-efficacy, there is no evidence that Induction 1 improved self-efficacy any more than the other inductions. There are several possible explanations for this finding. First, given the relatively short intervention, the fact that students in any induction did not experience dramatic gains in a construct as fundamental and multi-dimensional as self-efficacy is not surprising. Second, Induction 1 was the most complex in terms of cognitive and temporal "overhead" required for students to enact the experience; navigating and overcoming obstacles in a virtual world are more challenging tasks than the other inductions presented. We hypothesize that, had a longer time period been available for students to shift their focus from learning to enact Induction 1 to reflecting on the content of the experience, effects on self-efficacy would have been greater.

Finally, a third possible explanation for this finding is that, although all three inductions did target different aspects of motivation, these inductions were not the only component of the overall four-day intervention that was designed to influence students' motivation. In fact, the two-day mathematics lesson were also designed with best practices (including motivation) in mind. Given that the two-day math lesson was implemented with reasonably high fidelity, it may be that the mastery experiences afforded by the classroom lessons washed out any self-efficacy effects that the technologies provided. And when students thought about their confidence to do these types of problems in completing the self-efficacy items on the survey, they may have reflected more on their experiences in the classroom than in their respective technology experiences.

Related, recall that the three inductions also differed on the expense and technical sophistication required to create and implement them. Does the present finding about Induction 1 and self-efficacy suggest that use of virtual worlds is not worth the trouble and expense? Particularly when inculcating sophisticated knowledge and skills, a substantial body of research suggests that this is not the case (U.S. Department of Education, 2010; National Research Council, 2011; Author, in press). We interpret our results as indicating that this type of complex intervention with high cognitive overhead may require more instructional “dosage” than short duration provided in the present intervention. Thus, well-designed virtual worlds, which are expensive and technically demanding, can realize their power for engagement and learning only when a sufficient investment of classroom time is made.

Effects linked to students’ age. A second pattern that emerges from the complex results of our second research question is that the effects of each induction on students’ motivation were influenced by students’ age, as evidenced by the frequency of significant induction type by grade interactions. These grade-level interactions held while controlling for prior mathematics knowledge (VA-SOL scores), indicating that the differential impact of the inductions was developmental and not merely the result of differing mathematics ability. Because the structure of schooling for students in middle school (Grades 6-8) is different from that of elementary school students (Grade 5), and because students conceive of competence differently based on age (Dweck, 1986), these findings indicating differential impacts on students depending on their age are confirmatory of prior work and reinforce the importance for practitioners and policy makers of tailoring such interventions to students’ developmental level.

In addition, our results suggest that the impact of the abridged version of Brainology® on students’ implicit theory was greater for older students than it was for younger students. One possibility for this finding is that older students may be more attuned to the incremental message than younger students. Dweck (2002) argued that students’ conceptions of ability may not have an effect on their motivation and performance until 10-12 years old. Therefore, the incremental theory of ability message may have been more salient for these older students than it was for younger ones.

Absence of effects for student demographics. Finally, we did not find interactions between induction type and other student demographic variables such as free and reduced lunch, ethnicity, and gender. From a curricular perspective, this is a positive outcome indicating that, in contrast to many educational experiences, these types of intervention may narrow—not widen—troubling achievement gaps. That good design can produce motivational learning experiences effective across the full spectrum of students is very encouraging.

RQ3: Influences of Teacher-level Factors

Our third research question asked about impact of teacher-level factors on students' motivation, including credentialing in mathematics education, undergraduate major, years of experience, and teachers' beliefs. Based on the extant literature, we had hypothesized that these factors might influence students' motivation. However, teacher-level factors were not significant predictors of student outcomes. Viewing the intervention from a curricular perspective, this is a positive finding suggesting that our design and implementation ensured that all students received a roughly equivalent instructional experience.

With respect to the absence of a relationship between teachers' beliefs and student motivation, although there is good theoretical and empirical evidence to suggest that these variables could predict student outcomes, it is also true that linking teacher-level beliefs to student outcomes is not a clear and straight path (Holzberger, Phillipp, & Kunter, in press; Klassen et al., 2011). In fact, Klassen et al. (2011) noted that there is a lack of evidence that links teachers' self-efficacy to student outcomes, despite the commonly held belief by researchers that this relationship exists. Their review of the literature noted that correlations between teachers' self-efficacy and student achievement were low to modest. Our findings confirm this perspective.

One explanation for the absence of these effects may relate to a social desirability bias influencing teachers. We note that teachers' responses were generally quite positive on their motivation surveys, with relatively small variance. It may have been the case that some teachers were reluctant to admit they were not confident in being able to teach or manage a class effectively; similarly, some teachers might have been unwilling to admit that they saw little value in the goals of the present study. Artificially inflated teacher responses to the teacher motivational surveys may explain the lack of relationship between teacher and student motivation.

Another possibility is that the professional development that we created and implemented had the effect of eliminating much of the teacher-level variance and its effects on student outcomes. We specifically designed the professional development such that teachers emerged confident in their ability to successfully implement the two-day math lesson. We also communicated to teachers that there was considerably flexibility in their implementation of the math lesson, as long as a few basic implementation guidelines were followed. We hoped that such an empowerment-supportive way of training teachers would allow teachers to feel more autonomous and less controlled, thereby translating to better implemented curricula. It is possible that this approach (which did enable teachers to implement the two-day lesson with fidelity) also helps explain the absence of teacher-level effects on student motivation.

Limitations

There were several limitations to the present study that suggest caution in the interpretation of our results. First and foremost, as noted above, there was a very large amount of missing data – 53% of students were missing demographic, pre-, and/or posttest data – most of which occurred due to a miscommunication between the research team and the district relating to the student identification numbers that students were instructed to use at pre-test. Second, it is important to note that the length of the intervention was relatively short, both in terms of the technology-based motivational activities, the professional development, and the mathematics lesson. Although we were able to find some influence of the intervention on students' motivation, these effects were quite modest. Further, although a delayed posttest was administered, results were not interpretable; thus, we are not able to report whether or not the effects at posttest were sustained after the end of the intervention. Third, recall that the five-item math assessment had low reliability. Taken together, all of these results raise questions about any attempt to generalize our findings. Future studies – both additional large-scale studies of longer duration, as well as shorter-term studies that afford opportunities for more qualitative exploration - can attempt to address these limitations and continuing moving toward improving our understanding of the relationship between technology, motivation, and STEM learning.

Conclusion

Investigating along a developmental span the relationship between specific technology-based motivational activities and student interest in STEM careers is important, because much potential talent in science, technology, engineering, and mathematics is now lost. Our research interweaved alternative motivational activities with effective and authentic mathematics learning, in order to take initial steps toward developing insights about the added value of technology for building confidence in math and science capability, seeing one's abilities in STEM as able to improve over time, and developing a passion or sustained interest in becoming a scientist or engineer. Further, we studied the impacts of media with substantially different production costs, providing the basis for a cost-benefit analysis and for articulating contrasting conditions for success.

Our findings highlight the importance of tailoring motivational experiences to students' developmental level. Our results are also encouraging about developers' ability to create instructional interventions and professional development that can be effective when experienced by a wide range of students and teachers. Further research is needed to determine the degree, duration of, and type of instructional intervention necessary to substantially impact multi-dimensional, deep-rooted motivational constructs, such as self-efficacy.

References

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school though college*. Washington, DC: United States Department of Education.

Andrew, S. (1998). Self-efficacy as a predictor of academic performance in science. *Journal of Advanced Nursing*, 27, 596-603.

Archambault, I, Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from Grades 1 through 12. *Journal of Educational Psychology*, 102, 804-816.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents*. Greenwich, CT: Information Age Publishing (pp. 307-337).

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzales, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle years: IEA's third international mathematics and science study*. Boston: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Beghetto, R. A. (2007). Factors associated with middle and secondary students' perceived science competence. *Journal of Research in Science Teaching*, 44, 800–814.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and intervention. *Child Development*, 78, 246-263.

Blume, G. W., & Heckman, D. S. (1997). What do students know about algebra and functions? In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment* (pp. 225-277). Reston, VA: National Council of Teachers of Mathematics.

Bonous-Hammarth, M. (2000). Pathways to success: Affirming opportunities for science, mathematics, and engineering majors. *The Journal of Negro Education*, 69, 92-111.

Britner, S. L., & Pajares, F. (2001). Self-efficacy beliefs, race, and gender in middle school science. *Journal of Women and Minorities in Science and Engineering*, 7, 271-285.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44, 1187-1218.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Retrieved from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2×2 achievement goal framework. *Journal of Personality and Social Psychology*, 90, 666-679.

Donohoe, C., Topping, K., & Hannah, E. (2012). The impact of an online intervention (Brainology) on the mindset and resiliency of secondary school pupils: a preliminary mixed methods study. *Educational Psychology*, 32, 641-655.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. S. Eccles (Eds.), *The development of achievement motivation*. San Diego: Academic Press.

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Dweck, C. S., & Master, A. (2009). Self-theories and motivation: Students' beliefs about intelligence. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 123-140). New York: Routledge.

Eccles (Parsons), J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed), *Achievement and achievement motivation* (pp. 75-146). San Francisco: W. H. Freeman.

Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular activities and adolescent development. *Journal of Social Issues*, 59, 865-889.

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. Nicholls (Ed.), *Advances in motivation and achievement: The development of achievement motivation* (Vol. 3, pp. 283–331). Greenwich, CT: JAI Press.

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a model of social influence that explains minority integration into the scientific community. *Journal of Educational Psychology*, 103, 206-222.

Fox, J., & Bailenson, J. N. (2009). Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology*, 12, 1-25.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave MacMillan.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24, 645-662.

Good, C., Rattan, A, & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102, 700-717.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, 85, 541-553.

Gwilliam, L. R., & Betz, N. E. (2001). Validity of measures of math- and science-related self-efficacy for African Americans and European Americans. *Journal of Career Assessment,* 9, 261-281.

Hernandez, P. R., Schultz, P. W., Estrada, M., Woodcock, A., Chance, R. C. (2013). Sustaining optimal motivation: A longitudinal analysis of interventions to broaden participation of underrepresented students in STEM. *Journal of Educational Psychology*, 105, 89-107.

Hickey, D. T., Moore, A. L., & Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference?. *American Educational Research Journal*, 38, 611-652.

Holzberger, D., Philipp, A., & Kunter, M. (in press). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology*. DOI: 10.1037/a0032198

Hong, Y. Y., Chiu, C. Y., Dweck, C. S., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, 77, 588–599.

Hoyles, C., Noss, R., Kent, P., & Bakker, A. (2010). *Improving mathematics at work: The need for techno-mathematical literacies*. New York: Routledge.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102, 880.

Kaput J. (1999). Teaching and learn in a new algebra. In E. Fennema & T. Romberg (Eds.), *Mathematics classrooms that promote understanding* (p. 133-155). Mahwah, NJ: Erlbaum.

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 99-111.

Ketelhut, D. J., Nelson, B. C., Clarke, J. E., & Dede, C. (2010). A multi-user virtual environment for building and assessing higher order inquiry skills in science. *British Journal of Educational Technology*, 41(1), 56-68.

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390-419). New York: Simon & Schuster.

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, 23, 21- 43.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment*, 8, 139-162.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, 31, 356-362.

Lent, R. W., Brown, S. D., Sheu, H.-B., Schmidt, J., Brenner, B. R., Gloster, C., . . . Treistman, D. (2005). Social cognitive predictors of academic interest and goals in engineering: Utility for women and students at historically black universities. *Journal of Counseling Psychology*, 52, 84–92. doi:10.1037/0022-0167.52.1.84.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology*, 38, 424-430.

Lent, R.W., Lopez, A.M., Lopez, F.G., & Sheu, H. (2008). Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *Journal of Vocational Behavior*, 73, 52-62.

Lent, R.W., Paixão, M.P., da Silva, J.T., & Leitão, L.M. (2010). Predicting occupational interests and choice aspirations in Portuguese high school students: A test of social cognitive career theory. *Journal of Vocational Behavior, 76,* 244-251.

Lindquist, M. M. (Ed.). (1989). *Results from the fourth mathematics assessment of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), pp. 1198-1201.

Liu, M., Hsieh, P., Cho, Y., Schallert, D. L. (2006). Middle school students’ self-efficacy, attitudes, and achievement in a computer-enhanced problem-based learning environment. *Journal of Interactive Learning Research*, 17, 225-242.

Luzzo, D. A., Hasper, P., Albert, K. A., Bibby, M. A., & Martinelli, E. A., Jr. (1999). Effects of self-efficacy-enhancing interventions on the mathematics/science self-efficacy and career interests, goals, and actions of career undecided college students. *Journal of Counseling Psychology*, 46, 233-243.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, 95, 877–907. doi:10.1002/sce.20441

Moos, D. C., & Azevedo, R. (2008). Monitoring, planning, and self-efficacy during learning with hypermedia: The impact of conceptual scaffolds. *Computers in Human Behavior*, 24, 1686-1706.

Moos, D. C., & Marroquin, E. (2010). Multimedia, hypermedia, and hypertext: Motivation considered and reconsidered. *Computers in Human Behavior*, 26, 265-276.

Myers, C. B. (2011). Union status and faculty job satisfaction: Contemporary evidence from the 2004 National Study of Postsecondary Faculty. *The Review of Higher Education*, 34 (4), 657–684. doi:10.1353/rhe.2011.0028

National Academy of Sciences (2011). *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. Washington, DC: National Academies Press.

National Association of Education Progress, Question Tool. (2011). U.S. Department of Education. Retrieved from http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx?subject=mathematics

National Research Council. (2011). *Learning science through computer games and simulations*. Washington, DC: National Academy Press.

Oyserman, D., & Destin, M. (2010). Identity-based motivation: Implications for intervention. *The Counseling Psychologist*, 38, 1001-1043.

Rattan, A., Good, C., & Dweck, C. S. (2012). It’s ok-not everyone can be good at math: Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48, 731-737.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd edition). Thousand Oaks, CA: Sage Publications.

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior*, *24*(6), 2741–2756.

Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). *Facing the consequences: Using TIMMS for a closer look at U.S. mathematics and science education*. Dordrecht: Kluwer.

Squire, K. D. (2003). Video games in education. *International Journal of Intelligent Games & Simulation*, 2, 49-62.

Stipek, D., & Gralinski, J. H. (1996). Children's beliefs about intelligence and school performance. *Journal of Educational Psychology*, 88, 397-407.

U.S. Department of Education. (2010). *Transforming American education: Learning powered by technology* (National Educational Technology Plan 2010). Washington, DC: Office of Educational Technology, U.S. Department of Education. Retrieved from http://www.ed.gov/technology/netp-2010

Watt, H. M. G., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls: A motivational analysis of high school enrolments in Australia and the USA. *Equal Opportunities International*, 25, 642–659. doi:10.1108/02610150610719119

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal*, 37, 215-246.

Notes

[1] A delayed post-test was also administered, two months after the end of the intervention. However, due to large amounts of missing data, delayed post-test results were not easily interpretable and thus are not included in the present analysis.

[2] Student and teacher assessments also included additional items assessing several other motivational constructs. The inclusion of these extra items was exploratory, in that none of the technology-based activities were designed with these constructs in mind. In particular, students were administered a short assessment immediately after the conclusion of the Day 1 technology-based motivational activities that focused on several of these additional motivational constructs. In the present analysis, we report only on those student and teacher variables that were explicitly considered in the design of the inductions and that were specifically hypothesized to be related to the effectiveness of the intervention.

Table 1.
Student Demographic Information by Condition

Variable		*Induction 1*		*Induction 2*		*Induction 3*		*Total*	
Gender		*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
	Male	1373	51	1071	49	1516	50	3960	50
	Female	1308	49	1115	51	1517	50	3940	50
Ethnicity									
	Native American	11	<1	5	<1	7	<1	23	<1
	Asian	89	3	77	4	91	3	257	3
	African-American	691	26	516	24	647	21	1854	23
	Hispanic/Latino	260	10	194	9	202	7	656	8
	White	1500	56	1309	60	1938	64	4747	60
	Pacific Islander	1	<1	4	<1	4	<1	9	<1
	Multi-Race	129	5	81	4	144	5	354	4
Grade									
	5	768	29	523	24	845	28	2136	27
	6	877	33	370	17	515	17	1762	22
	7	572	21	615	28	898	30	2085	26
	8	464	17	678	31	775	26	1917	24
ELL		125	5	81	4	83	3	289	4

Table 2.
Descriptive Statistics on Student Motivation and Learning Variables

Variable	Pretest									Posttest								
	Induction 1		Induction 2		Induction 3		Total			Induction 1		Induction 2		Induction 3		Total		
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*n*
VA-SOL	498	75	491	80	497	78	496	78	7900	-	-	-	-	-	-	-	-	-
Math learning	0.60	0.24	0.61	0.23	0.60	0.24	0.60	0.24	7900	0.68	0.25	0.70	0.24	0.71	0.24	0.70	0.24	6983
Self-efficacy	4.59	0.99	4.49	1.02	4.53	1.00	4.54	1.00	7900	4.60	1.07	4.54	1.08	4.51	1.09	4.55	1.08	7045
Implicit theory of math ability	4.26	1.04	4.17	1.03	4.24	1.03	4.22	1.03	7900	4.09	1.07	4.27	1.08	4.14	1.08	4.16	1.08	7090
Value	4.33	1.00	4.16	1.07	4.23	1.04	4.24	1.04	7900	4.28	1.12	4.14	1.15	4.14	1.15	4.19	1.14	7063

Table 3.
Motivational Measures

	Construct	*Alpha*	*Measure*	*Sample Question (all on a 6 point scale)*
STUDENT MEASURES	Self-Efficacy (n = 13)	0.93, 0.95	General Math Self-Efficacy (n = 4)	How confident are you that you can master the math skills that will be taught this year?
			Algebraic Reasoning Self-Efficacy (n = 5)	If you are given 5 numbers in a sequence, how confident are you that you can figure out the pattern and get the next number in the sequence right?
			Math Performance Self-Efficacy (n = 4)	How confident are you that you can do well on standardized tests in math?
	Implicit Theory of Math Ability (n = 6)	0.77, 0.79	Fixed View of Math Ability (n = 3)	My math ability is something about me that can't be changed very much.
			Incremental View of Math Ability (n = 3)	No matter who I am, I can change my math abilities a lot.
	Value (n = 6)	0.83, 0.87	Interest Value (n = 3)	How much do you like math?
			Utility Value (n = 2)	In general, how useful is what you learn in math?
			Attainment Value (n = 1)	For me, how important is being good at math?
TEACHER MEASURES	Self-Efficacy for Instruction and Student Engagement (n = 22)	0.96	Self-Efficacy for Student Engagement (n = 4)	How confident are you that you can motivate students who show low interest in math class?
			Self-Efficacy for Classroom Management (n = 4)	How confident are you that you can calm a student who is disruptive and noisy?
			Self-Efficacy for Instructional Strategies (n = 4)	How confident are you that you can use a variety of assessment strategies?
			Self-Efficacy for Math Inquiry Teaching (n = 6)	How confident are you that you can use computer technologies to communicate with your students?
			Self-Efficacy for Instructional Methods (n = 4)	How confident are you that you can teach well even if you are told to use instructional methods that would not be your choice?
	Self-Efficacy for Technology Use (n = 7)	0.89		How confident are you that you can facilitate a whole-class discussion?
	Math Self-Efficacy (n = 12)	0.92		How confident are you that you can successfully determine the amount of sales tax on a clothing purchase?
	Implicit Theory of Math Ability (n = 6)	0.86	Fixed View about Students' Abilities in Math (n = 3)	Students come in to math with a certain level of math ability, and it is hard to change that.
			Incremental View About Students' Abilities in Math (n = 3)	Even if students don't initially possess a certain "knack" for math they can develop their math ability.

Table 4.
Parameter Estimates for Student Outcomes

	Posttest Math Learning			Posttest Self-Efficacy		
Fixed Effects	Coefficient	SE	z	Coefficient	SE	z
Intercept	0.67	0.02	41.82***	4.67	0.04	107.55***
Student-level						
VASOL	0	0	11.34***	0	0	1.49
Pretest math learning	0.20	0.01	16.42***	0.14	0.04	3.69***
Pretest self-efficacy	0.02	0	5.87***	0.70	0.01	62.80***
Pretest implicit theory of math ability	0	0	0.44	0.05	0.01	6.20***
Pretest value	0.01	0	3.99***	0.15	0.01	14.11***
ELL status	-0.01	0.01	-0.70	-0.07	0.04	-1.70
Grade	0	0.01	0.08	-0.05	0.02	-2.87**
Gender (Male)	-0.02	0.01	-3.61***	0	0.02	0.05
Ethnicity	0	0	0.99	-0.01	0.01	-1.89τ
Teacher-level						
Induction 2	0	0.02	0.16	-0.01	0.04	-0.18
Induction 3	-0.01	0.01	-0.83	-0.04	0.03	-1.27
Self-efficacy for student engagement and instruction	0.02	0.01	2.18*	0.02	0.02	1.00
Self-efficacy for technology use	-0.02	0.01	-3.07**	-0.02	0.01	-1.58
Math self-efficacy	0	0.01	-0.06	0	0.01	-0.39
Implicit theory of math ability	0	0.01	-0.06	-0.01	0.01	-0.83
School-level						
% free/reduced lunch	-0.11	0.03	-3.90***	-0.06	0.06	-0.92
Cross-level interactions						
Induction 2 by Grade	0.01	0.01	1.52	0.03	0.02	1.52
Induction 3 by Grade	0.02	0.01	2.49*	0	0.02	0.17
Induction 2 by VASOL	0	0	0.62	0	0	0.25
Induction 3 by VASOL	0	0	0.23	0	0	0.94
Random Effects	Estimate	SE		Estimate	SE	
Level-1 residual variance	0.21	0		0.66	0.01	
Level-2 residual variance	0.05	0		0.07	0.01	
Level-3 residual variance	0.01	0.01		0	0	

	Posttest Implicit Theory of Math Ability			Posttest Value		
Fixed Effects	Coefficient	SE	z	Coefficient	SE	z
Intercept	4.16	0.05	79.50***	4.28	0.04	96.23***
Student-level						
VASOL	0	0	-0.04	0	0	1.74
Pretest math learning	0.04	0.05	0.86	0	0.04	-0.01
Pretest self-efficacy	0.09	0.01	6.32***	0.09	0.01	7.79***
Pretest implicit theory of math ability	0.60	0.01	56.85***	0.02	0.01	2.76**
Pretest value	0.08	0.01	6.21***	0.83	0.01	79.77***
ELL status	-0.04	0.05	-0.76	0.07	0.04	1.65
Grade	-0.07	0.02	-3.55***	0	0.02	-0.26
Gender (Male)	-0.05	0.02	-2.51*	0	0.02	-0.03
Ethnicity	0	0.01	0.32	-0.02	0.01	-2.16*
Teacher-level						
Induction 2	0.09	0.05	2.07*	0.02	0.04	0.43
Induction 3	0.05	0.04	1.17	0.01	0.04	0.26
Self-efficacy for student engagement and instruction	0.04	0.02	1.84	0.01	0.02	0.70
Self-efficacy for technology use	-0.01	0.02	-0.52	-0.01	0.01	-0.56
Math self-efficacy	-0.03	0.01	-1.76	0.01	0.01	0.55
Implicit theory of math ability	-0.03	0.02	-1.88	0.02	0.01	1.39
School-level						
% free/reduced lunch	0	0.08	0	0.03	0.07	0.38
Cross-level interactions						
Induction 2 by Grade	0.12	0.03	4.57***	-0.01	0.02	-0.64
Induction 3 by Grade	0.03	0.02	1.10	-0.04	0.02	-2.10*
Induction 2 by VASOL	0	0	2.37*	0	0	-1.82
Induction 3 by VASOL	0	0	0.89	0	0	-1.34
Random Effects	Estimate	SE		Estimate	SE	
Level-1 residual variance	0.81	0.01		0.66	0.01	
Level-2 residual variance	0.07	0.02		0.08	0.01	
Level-3 residual variance	0.02	0.03		0.03	0.02	

τ *p* < .06, **p* < .05; ** *p* < .01, *** *p* < .001

Appendix

Table A-1.
Differences Between Students Included and Excluded

Variable		*Included*		*Excluded*		*Total*	
Gender***		*n*	*%*	*n*	*%*	*n*	*%*
	Male	3960	50	4682	52	8642	51
	Female	3940	50	4242	47	8182	48
Ethnicity***							
	Nat. American	23	<1	36	<1	59	<1
	Asian	257	3	297	3	554	3
	African-American	1854	23	2782	31	4636	27
	Hispanic/Latino	656	8	1007	11	1663	10
	White	4747	60	4464	50	9211	55
	Pacific Islander	9	<1	11	<1	20	<1
	Multi-Race	354	4	326	4	680	4
Grade***							
	5	2136	27	2090	23	4226	25
	6	1762	22	2494	28	4256	25
	7	2085	26	2127	24	4212	25
	8	1917	24	2212	25	4129	24
ELL***							
	Non-ELL	7611	96	8217	92	15828	94
	ELL	289	4	538	6	827	5
Induction***							
	1	2681	34	2470	28	5151	31
	2	2186	28	3262	36	5448	32
	3	3033	38	2537	28	5570	33

*** Differences between included and excluded students significant at the $p < .001$ level

Table A-2.
Fidelity Measures and Results

Measure	Item	% Respondents					
		Poor	*Not Very Good*	*Neither Bad nor Good*	*Very Good*	*Excellent*	*Blank*
Teacher Evaluation of Professional Development *(administered after PD)*	Compared to other teacher professional development experiences you have had in the past 5 years, how would you rate the overall quality of the one-day professional development program that you just completed?	6	17	37	26	3	11
		Not at all	*Somewhat*	*Moderately*	*Very*	*Completely*	*Blank*
	As a result of your experiences today, how PREPARED do you feel to teach the two-day mathematics lessons?	1	12	31	35	10	11
	How CONFIDENT are you that you will be able to successfully implement the two-day mathematics lessons?	1	8	21	42	16	12
	How SIMILAR is the two-day mathematics lesson to a typical mathematics lesson that you have taught this year?	18	28	26	15	1	11
	How CHALLENGED do you anticipate your students will be with the CONTENT of the two-day mathematics lessons?	2	6	29	37	16	11
		Definitely will NOT Enjoy	*Probably Will NOT Enjoy*	*Not Sure*	*Probably WILL Enjoy*	*Definitely WILL Enjoy*	*Blank*
	How do you anticipate your students will REACT to the two-day lessons?	2	12	27	41	6	11
Teacher Self-reported Fidelity of Implementation *(administered after two-day math lesson)*[a]	*Considering the Condensed Lesson Plan for the two-day math lesson that was included in your curriculum manual and is provided here,*	*Not at All*	*Not Very Closely*	*Somewhat Closely*	*Very Closely*	*Exactly*	*Blank*
	Please indicate how closely you followed the list of activities in the TO	<1	2	23	64	11	<1

DO column on the left side.						
Please indicate how closely you followed the list of activities in the TO SAY column on the right side.	1	4	36	51	9	<1
Compared to the TIMES recommended in the Condensed Lesson Plan….	*Not at All*	*Not Very Closely*	*Somewhat*	*Very Closely*	*Exactly*	*Blank*
The amount of time that I spent on the DEFINITION OF TASK sections of the lesson (colored BLUE in the Condensed Lesson Plan)	2	12	55	21	3	8
The amount of time that I spent on the EXPLORATORY PROBLEM SOLVING section of the lesson (colored LIGHT GREEN in the Condensed Lesson Plan)	2	9	52	25	5	8
The amount of time that I spent on the REPRESENTATION OF DATA section of the lesson (colored ORANGE in the Condensed Lesson Plan)	1	7	36	22	5	29
The amount of time that I spent on the PATTERN IDENTIFICATION, DESCRIPTION, AND GENERALIZATION section of the lesson (colored RED in the Condensed Lesson Plan)	1	11	45	27	7	8

[a] Note. For 6th, 7th, and 8th grade teachers who taught more than one class period of the intervention, teachers answered the fidelity of implementation questions separately for each period.